EXPERT SYSTEM VERIFICATION AND VALIDATION STUDY

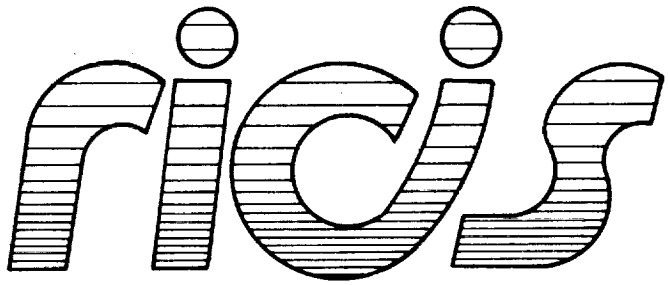
Delivery 1 - Survey and Interview Questions

International Business Machines Corporation

June 22, 1990

Cooperative Agreement NCC 9-16
Research Activity No. AI.16

NASA Johnson Space Center
Information Systems Directorate
Information Technology Division



Research Institute for Computing and Information Systems
University of Houston - Clear Lake

T·E·C·H·N·I·C·A·L R·E·P·O·R·T

The RICIS Concept

The University of Houston-Clear Lake established the Research Institute for Computing and Information systems in 1986 to encourage NASA Johnson Space Center and local industry to actively support research in the computing and information sciences. As part of this endeavor, UH-Clear Lake proposed a partnership with JSC to jointly define and manage an integrated program of research in advanced data processing technology needed for JSC's main missions, including administrative, engineering and science responsibilities. JSC agreed and entered into a three-year cooperative agreement with UH-Clear Lake beginning in May, 1986, to jointly plan and execute such research through RICIS. Additionally, under Cooperative Agreement NCC 9-16, computing and educational facilities are shared by the two institutions to conduct the research.

The mission of RICIS is to conduct, coordinate and disseminate research on computing and information systems among researchers, sponsors and users from UH-Clear Lake, NASA/JSC, and other research organizations. Within UH-Clear Lake, the mission is being implemented through interdisciplinary involvement of faculty and students from each of the four schools: Business, Education, Human Sciences and Humanities, and Natural and Applied Sciences.

Other research organizations are involved via the "gateway" concept. UH-Clear Lake establishes relationships with other universities and research organizations, having common research interests, to provide additional sources of expertise to conduct needed research.

A major role of RICIS is to find the best match of sponsors, researchers and research objectives to advance knowledge in the computing and information sciences. Working jointly with NASA/JSC, RICIS advises on research needs, recommends principals for conducting the research, provides technical and administrative support to coordinate the research, and integrates technical results into the cooperative goals of UH-Clear Lake and NASA/JSC.

EXPERT SYSTEM VERIFICATION AND VALIDATION STUDY

Delivery 1 - Survey and Interview Questions

Preface

This research was conducted under auspices of the Research Institute for Computing and Information Systems by the International Business Machines Corporation. Dr. Terry Feagin and Dr. T. F. Leibfried served as RICIS research representatives.

Funding has been provided by Information Technology Division, Information Systems Directorate, NASA/JSC through Cooperative Agreement NCC 9-16 between NASA Johnson Space Center and the University of Houston-Clear Lake. The NASA technical monitor for this activity was Chris Culbert, of the Software Technology Branch, Information Technology Division, Information Technology Directorate, NASA/JSC.

Expert Systems Evaluation Questionnaire (Developer)

By filling out this NASA funded questionnaire, you can help define the state-of-the-practice in the formal evaluation of Expert Systems on current NASA and industry applications. The information that you provide will be merged with the information from all other surveyed projects for the purpose of recommending future research and development activities. Individual responses are used solely as input to this information merging process.

Expert System applications are becoming more prevalent in fields where proper functioning is essential, such as the medical, financial, and aerospace industries. It is widely claimed that Expert Systems are not as rigorously evaluated as traditional software because of unique, unresolved evaluation issues. To ensure the continued and safe deployment of Expert Systems into critical areas, adequate evaluation techniques which address these issues must be developed and performed.

The answers to this questionnaire, together with follow-up interviews, will provide realistic answers to the following questions:

- How much evaluation is being performed?
- What evaluation techniques are in use?
- What, if any, are the unique issues in evaluating Expert Systems?

Instructions

The following questions concern your experiences with an Expert System, either as a developer or as the manager of the development effort. Feel free to indicate your answers in any way you like. Some of the choices on the multiple choice questions have places to fill in additional information; please indicate the choice and include the additional information, if possible.

This survey task must be performed within a relatively short time period. If possible, please return completed questionnaires within one week of receipt to:

Keith Kelley
MC 6272A
IBM Federal Solutions Division
3700 Bay Area Blvd.
Houston, Tx. 77058-1199

If you have any questions regarding this questionnaire, please contact Keith at (713) 282-7303. Each participating project may request a copy of the final survey report from Keith.

Questions

1. What is the name of the Expert System you were/are involved with?

   _____

2. Were you a developer of the Expert System or the manager of the development organization?

   a. Developer of Expert System
   b. Manager of Expert System development organization

3. The responses that you provide in this questionnaire may indicate that further discussion is required for us to understand the issues that you encountered during the evaluation process. Would you be available, at your convenience, to discuss the evaluation process in more detail? Interviews will be fairly short one-on-one meetings either in person or by telephone.

   a. No, I am not available for an interview.

   b. Yes, I am available.

      Name _____

      Phone _____

4. What field does the problem belong to?

   a. Aerospace
   b. Financial
   c. Information Systems
   d. Hardware
   e. Manufacturing
   f. Marketing
   g. Medical
   h. Personnel
   i. Research
   j. Service
   k. Software
   l. Other _____

5. Which of the following items best describes what the Expert System does? Please indicate primary purpose with a '*' and check all other applicable purposes (if any).

   a. Design - Configuring objects under constraints
   b. Repair - Executing plans to administer prescribed remedies
   c. Control - Governing overall system behavior
   d. Planning - Designing actions
   e. Diagnosis - Inferring system malfunctions from observables
   f. Debugging - Prescribing remedies for malfunctions
   g. Prediction - Inferring likely consequences of given situations
   h. Monitoring - Comparing observations to expected outcomes
   i. Instruction - Diagnosing, debugging, and repairing behavior
   j. Interpretation - Inferring situation descriptions from sensor
   k. Classification - Categorizing objects by properties data

6. Some problems require the use of certainty factors (also called probabilities, or "fuzzy logic") in their processing. Facts which contain certainty factors have the form: "if a is true, then there is an x% chance that b is true." Does the Expert System include certainty factors?

a. Yes
b. No
c. I don't know

7. How much of the total problem space is the Expert System expected to address? I.e., if the Expert System is supposed to be able to diagnose 100 malfunctions, but the total number of malfunctions is known to be 200, the Expert System is expected to address 50% of the problem space.

a. 100%
b. > 99%
c. 95% to 99%
d. 90% to 95%
e. 80% to 90%
f. 60% to 80%
g. 40% to 60%
h. Other _____%
i. I don't know

8. What is *your estimate* of the percentage of the problem space that the Expert System actually covers?

a. Same as expected
b. 100%
c. > 99%
d. 95% to 99%
e. 90% to 95%
f. 80% to 90%
g. 60% to 80%
h. 40% to 60%
i. Other _____%
j. I don't know

9. When developing an Expert System, the person who provides the knowledge that is to be captured in the system is called the *expert*. For that part of the problem space addressed by the Expert System, how often is the expert(s) expected to give the correct answer? I.e., referring to question 7, how often does the expert(s) identify the correct malfunction out of the 100 addressed malfunctions?

a. "Correct" defined by expert
b. > 99%
c. 95% to 99%
d. 90% to 95%
e. 80% to 90%
f. 60% to 80%
g. 40% to 60%
h. Other _____%
i. I don't know

10. For that part of the problem space addressed by the Expert System, how often is the Expert System expected to provide the correct answer?

a. 100%
b. > 99%
c. 95% to 99%
d. 90% to 95%
e. 80% to 90%
f. 60% to 80%
g. 40% to 60%
h. Other _____%
i. I don't know

11. What is *your estimate* of the percentage of the time that the Eepert System provides the correct answer for that part of the problem space addressed by the Expert System?

a. 100%
b. > 99%
c. 95% to 99%
d. 90% to 95%
e. 80% to 90%
f. 60% to 80%
g. 40% to 60%
h. Other _____%
i. I don't know

12. What was the basis for determining how the system was to behave?

a. A pre-existing document _____
b. A requirements document completed as part of development.
c. Some other developed document _____
d. A prototype of the system
e. Expert consultation
f. Other _____

13. Was there more than one expert consulted during the development of the system?

a. System was developed by expert
b. Single expert
c. Multiple experts with lead
d. Committee of experts
e. Other _____

14. How much interaction was there between the expert(s) and the development team?

a. Constant
b. Frequent
c. Regular
d. Occasional
e. None

15. Was the developer(s) part of the user organization?

a. Yes
b. No
c. User organization participated in development

16. Please indicate which development model was used for developing the Expert System.

a. Traditional waterfall life-cycle
b. Requirements gathered before development of a prototype. A second requirements activity preceded Design, Implementation, and Test.
c. Repetition of the Requirements, Design, Rule Generation, and Prototyping phases until production system (final prototype) was developed.
d. No effort was made to follow a particular model.
e. Other _____

17. What percentage of the total development effort was dedicated to the each of the three parts of the Expert System?

    a. Information Structures _____% (Declarative part which represents the knowledge of the Expert System.)

    b. Inference Engine _____% (Processes the knowledge base to infer a set of output facts from a set of input facts. If an Expert System shell was used, this value should be 0%.)

    c. Traditional Code _____% (Used to supplement the inference process (e.g., interfacing the inference engine to a device or user, performing arithmetic calculations, etc.).)

18. What was the primary language/tool for each part of the Expert System?

    a. Knowledge Base _____

    b. Inference Engine _____

    c. Traditional Code _____

19. How hard was it to develop the original concept of what the system was supposed to do?

    a. Trivial
    b. Easy
    c. Medium
    d. Hard
    e. Impossible

20. Aside from any difficulties in developing the original concept, how hard was it to express the behavior (through the knowledge base)?

    a. Trivial
    b. Easy
    c. Medium
    d. Hard
    e. Impossible

21. When changes were made to the knowledge, how often did some unexpected change occur?

    a. Never
    b. Occasionally
    c. Frequently
    d. Usually
    e. Always

22. How were changes the the Expert System distributed to the users?

    a. Developers made changes to users' system.

    b. Tested system distributed to the users.

    c. Other _____

23. If the Expert System is rule-based, how many rules are contained in the Knowledge Base (KB)? _____

If the Expert System is not rule-based, please give an approximate size of the knowledge base _____.

The answer just given is:

a. The actual effort
b. A very close approximation
c. A guess

24. How much effort was expended in developing the system? (Please take a guess if you don't know) _____ person/months.

The indicated effort is:

a. The actual effort
b. A very close approximation
c. A guess

25. Were any evaluation activities performed on the system while it was being developed? (indicate any that apply)

a. No evaluation was performed
b. Desk checking
c. Formal inspections
d. Checked by expert(s)
e. Compared with documented behavior
f. Structural testing (e.g. cover all rules)
g. Other _____

26. What evaluation activities were performed on the executing system after development was completed? (indicate any that apply)

a. No evaluation was performed
b. Checked by expert(s)
c. Compared with documented behavior
d. User acceptance
e. Other _____

27. During evaluation, the results from executing the system were compared with:

a. Requirements document
b. System prototype
c. Single expert
d. Majority opinion of experts
e. Other _____

28. What was the level of agreement among the experts concerning the correctness of the system? That is, do the experts agree on the correctness of the results of the system? Please note that this does not mean that the experts agree with the system, but rather, that they agree with each other about the results of the system.

a. Always agree
b. Agree _____% of the time.
c. A single expert was involved

29. How hard was the evaluation effort to perform?

a. Trivial
b. Easy
c. Medium
d. Hard
e. Impossible

30. How much effort was expended by the development organization in evaluating the correctness of the Expert System? (Please take a guess if you don't know) _____ person/months.

The indicated effort is:

a. The actual effort
b. A very close approximation
c. A guess

31. What is the worst thing that can happen if the Expert System gives the wrong answer?

a. Someone gets hurt
b. Loss of "mission"
c. Nuisance (correct answer derived some other way)
d. Work-around must be used
e. Nothing
f. Can't tell the answer is wrong
g. Other _____

32. How does the number of errors that the users encounter compare with the number of errors they encounter with other systems which are not Expert Systems?

a. Significantly more errors
b. More errors
c. About the same number of errors
d. Fewer errors
e. Significantly fewer errors
f. I don't know

Expert Systems Evaluation Questionnaire (User)

By filling out this NASA funded questionnaire, you can help define the state-of-the-practice in the formal evaluation of Expert Systems on current NASA and industry applications. The information that you provide will be merged with the information from all other surveyed projects for the purpose of recommending future research and development activities. Individual responses are used solely as input to this information merging process.

Expert System applications are becoming more prevalent in fields where proper functioning is essential, such as the medical, financial, and aerospace industries. It is widely claimed that Expert Systems are not as rigorously evaluated as traditional software because of unique, unresolved evaluation issues. To ensure the continued and safe deployment of Expert Systems into critical areas, adequate evaluation techniques which address these issues must be developed and performed.

The answers to this questionnaire, together with follow-up interviews, will provide realistic answers to the following questions:

- How much evaluation is being performed?
- What evaluation techniques are in use?
- What, if any, are the unique issues in evaluating Expert Systems?

Instructions

The following questions concern your experiences with an Expert System, either as a user or as the manager of a department that uses Expert System. Feel free to indicate your answers in any way you like. Some of the choices on the multiple choice questions have places to fill in additional information; please indicate the choice and include the additional information, if possible.

This survey task must be performed within a relatively short time period. If possible, please return completed questionnaires within one week of receipt to:

Keith Kelley
MC 6272A
IBM Federal Solutions Division
3700 Bay Area Blvd.
Houston, Tx. 77058-1199

If you have any questions regarding this questionnaire, please contact Keith at (713) 282-7303. Each participating project may request a copy of the final survey report from Keith.

Questions

1. What is the name of the Expert System you were/are involved with?

   _____

2. Are you a user of the Expert System or the manager of a department which uses the Expert System?

   a. User of the Expert System
   b. Manager of a department using the Expert System

3. The responses that you provide in this questionnaire may indicate that further discussion is required for us to understand the issues that you encountered during the evaluation process. Would you be available, at your convenience, to discuss the evaluation process in more detail? Interviews will be fairly short one-on-one meetings either in person or by telephone.

   a. No, I am not available for an interview.

   b. Yes, I am available.

      Name _____

      Phone _____

4. What field does the problem belong to?

   a. Aerospace
   b. Financial
   c. Information Systems
   d. Hardware
   e. Manufacturing
   f. Marketing
   g. Medical
   h. Personnel
   i. Research
   j. Service
   k. Software
   l. Other _____

5. Which of the following items best describes what the Expert System does? Please indicate primary purpose with a '*' and check all other applicable purposes (if any).

   a. Design - Configuring objects under constraints
   b. Repair - Executing plans to administer prescribed remedies
   c. Control - Governing overall system behavior
   d. Planning - Designing actions
   e. Diagnosis - Inferring system malfunctions from observables
   f. Debugging - Prescribing remedies for malfunctions
   g. Prediction - Inferring likely consequences of given situations
   h. Monitoring - Comparing observations to expected outcomes
   i. Instruction - Diagnosing, debugging, and repairing behavior
   j. Interpretation - Inferring situation descriptions from sensor data
   k. Classification - Categorizing objects by properties

6. Some problems require the use of certainty factors (also called probabilities, or "fuzzy logic") in their processing. Facts which contain certainty factors have the form: "if a is true, then there is an x% chance that b is true." Does the Expert System include certainty factors?

a. Yes
b. No
c. I don't know

7. How much of the total problem space is the Expert System expected to address? I.e., if the Expert System is supposed to be able to diagnose 100 malfunctions, but the total number of malfunctions is known to be 200, the Expert System is expected to address 50% of the problem space.

a. 100%
b. > 99%
c. 95% to 99%
d. 90% to 95%
e. 80% to 90%
f. 60% to 80%
g. 40% to 60%
h. Other _____%
i. I don't know

8. What is *your estimate* of the percentage of the problem space that the Expert System actually covers?

a. Same as expected
b. 100%
c. > 99%
d. 95% to 99%
e. 90% to 95%
f. 80% to 90%
g. 60% to 80%
h. 40% to 60%
i. Other _____%
j. I don't know

9. When developing an Expert System, the person who provides the knowledge that is to be captured in the system is called the *expert*. For that part of the problem space addressed by the Expert System, how often is the expert(s) expected to give the correct answer? I.e., referring to question 7, how often does the expert(s) identify the correct malfunction out of the 100 addressed malfunctions?

a. "Correct" defined by expert
b. > 99%
c. 95% to 99%
d. 90% to 95%
e. 80% to 90%
f. 60% to 80%
g. 40% to 60%
h. Other _____%
i. I don't know

10. For that part of the problem space addressed by the Expert System, how often is the Expert System expected to provide the correct answer?

a. 100%
b. > 99%
c. 95% to 99%
d. 90% to 95%
e. 80% to 90%
f. 60% to 80%
g. 40% to 60%
h. Other _____%
i. I don't know

11. What is *your estimate* of the percentage of the time that the Expert System provides the correct answer for that part of the problem space addressed by the Expert System?

a. Same as expected
b. 100%
c. > 99%
d. 95% to 99%
e. 90% to 95%
f. 80% to 90%
g. 60% to 80%
h. 40% to 60%
i. Other _____%
j. I don't know

12. Was the expert(s) a member of the user organization?

a. Yes
b. No
c. User organization provided some expertise

13. Was the developer(s) of the Expert System part of the user organization?

a. Yes
b. No
c. User organization participated in development

14. Why do you believe the results that the system gives?

a. Expert says it is correct
b. Participated in evaluation
c. Someone I trust did evaluation
d. Personal use and checking
e. User acceptance
f. I don't trust the results
g. Other _____

15. What is the worst thing that can happen if the Expert System gives the wrong answer?

a. Someone gets hurt
b. Loss of "mission"
c. Nuisance (correct answer derived some other way)
d. Workaround must be used
e. Nothing
f. Can't tell answer is wrong
g. Other _____

16. How does the number of errors that the users encounter compare with the number of errors they encounter with other systems which are not Expert Systems?

a. Significantly more errors
b. More errors
c. About the same number of errors
d. Fewer errors
e. Significantly fewer errors
f. No errors encountered
g. I don't know

If you were not involved with evaluating the Expert System, please leave the remaining questions unanswered.

17. What evaluation activities were performed on the executing system? (indicate any that apply)

a. No evaluation was performed
b. Checked by expert(s)
c. Compared with documented behavior
d. User acceptance
e. Other _____

18. During evaluation, the results from executing the system were compared with:

a. Requirements document
b. System prototype
c. Single expert
d. Majority opinion of experts
e. Other _____

19. What was the level of agreement among the experts concerning the correctness of the system? That is, do the experts agree on the correctness of the results of the system? Please note that this does not mean that the experts agree with the system, but rather, that they agree with each other about the results of the system.

a. Always agree
b. Agree _____% of the time.
c. A single expert was involved

20. How hard was the evaluation effort to perform?

a. Trivial
b. Easy
c. Medium
d. Hard
e. Impossible

21. How much effort was expended by the user group in evaluating the correctness of the Expert System? (please take a guess if you don't know) _____ person/months.

The indicated effort is:

a. The actual effort
b. A very close approximation
c. A guess